Version ICL/3

EGT3
ENGINEERING TRIPOS PART IIB

Friday 8 May 2026   9.30 to 11.10

Module 4F3

AN OPTIMISATION BASED APPROACH TO CONTROL

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

Write your candidate number ***not*** *your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Attachment: 4F3 data sheet (two pages).
Engineering Data Book

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 (a) Discuss the connection between Dynamic Programming and the Hamilton-Jacobi-Bellman (HJB) equation. Discuss also one advantage and one disadvantage of using the HJB equation to solve optimal control problems. [15%]

(b) Consider the system

$$x_{k+1} = \lambda x_k + u_k \tag{1}$$

where $x_k \in \mathbb{R}$, $u_k \in \mathbb{R}$ for $k = 0, \ldots, h-1$, $x_h \in \mathbb{R}$, and $\lambda \in \mathbb{R}$ is a constant. Consider also the cost

$$\gamma x_h^2 + \sum_{k=0}^{h-1} x_k^2 + u_k^2$$

where $\gamma > 0$ is a constant.

(i) Show using dynamic programming that the minimum value of this cost over all control inputs u_k, for a given initial condition x_0, is a quadratic function of the initial condition. [25%]

(ii) Consider the system (1) with $\lambda = 1$. If the minimum cost for a horizon of length $h = 4$ is $2x_0^2$, find the minimum cost in terms of the initial condition when $h = 6$. [10%]

(c) For system (1) consider the problem of minimizing the cost

$$\sum_{k=0}^{\infty} x_k^2 + u_k^2 \tag{2}$$

over the control inputs $u_k \in \mathbb{R}$. Discuss whether control inputs such that x_k becomes zero in finite time, can lead to a lower cost from that obtained from the best linear controller of the form $u_k = Kx_k$ where K is a constant. Find also an expression for the minimum value of the cost in (2) for a given initial condition x_0. [25%]

(d) Consider the optimal control problem in part (c), but the control input u_k can now take only discrete values in the set $\{-2, -1, 0, 1, 2\}$. Find the optimal control inputs and the minimum cost for initial condition $x_0 = 2$, with $\lambda = 1$ in system (1). [25%]

2 (a) Explain what is meant by the $\mathcal{H}_2$ optimal control problem. Also discuss in what ways $\mathcal{H}_2$ optimal control provides a generalization to the Linear Quadratic Regulator problem. [10%]

(b) Consider a stable linear system with state space representation $\dot{x} = Ax + Bw,\ y = Cx$. We denote the transfer function from w to y by G.

(i) Show that if there exists a constant matrix $M > 0$ such that $V(x) = x^T M x$ satisfies

$$\frac{dV}{dt} \leq \gamma^2 w^T w - y^T y$$

for all signals w and corresponding outputs y, such that $\int_0^\infty w^T w dt < \infty$, where $\gamma > 0$ is a constant, then $\|G\|_\infty \leq \gamma$. [15%]

(ii) Use this to derive a Riccati equation through which one can verify if $\|G\|_\infty \leq \gamma$. [25%]

(c) Consider the system

$$\begin{aligned} \dot{x} &= -0.2x + u + w_1 \\ y &= x + w_2 \\ z_1 &= x, \quad z_2 = u \end{aligned}$$

where $x(t) \in \mathbb{R}$ is the state of the system, $w_1(t), w_2(t)$ are external disturbances, and $u(t)$ is the control input. We also denote $w = [\ w_1 \quad w_2\ ]^T$, and $z = [\ z_1 \quad z_2\ ]^T$. A controller with transfer function $K(s)$ is implemented such that $\bar{u}(s) = K(s)\bar{y}(s)$, where $\bar{u}(s), \bar{y}(s)$ are the Laplace transforms of the signals u, y respectively.

(i) Find a controller $K(s)$ that will ensure the $\mathcal{H}_\infty$ norm of the transfer function from w to z is less than or equal to 2. [30%]

(ii) Consider a controller of the form $u = ky$ where $k \leq 0$ is a constant. Find an expression for the $\mathcal{H}_\infty$ norm of the transfer function from w to x. Discuss also if this $\mathcal{H}_\infty$ norm can be made arbitrarily small via appropriate choice of k. [20%]

3 Consider the discrete-time linear system

$$x(k+1) = Ax(k) + Bu(k),$$

with $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^m$, and assume full state measurement is available.

At time k, a predictive controller computes an input sequence

$$u = \{u_0, u_1, \ldots, u_{N-1}\}$$

and associated predicted states $\{x_i\}_{i=0}^{N}$ satisfying

$$x_0 = x(k), \qquad x_{i+1} = Ax_i + Bu_i, \quad i = 0, 1, \ldots, N-1,$$

to minimise the finite-horizon cost

$$V(x(k), u) = x_N^\top P x_N + \sum_{i=0}^{N-1} \left(x_i^\top Q x_i + u_i^\top R u_i \right),$$

where $Q > 0$, $R > 0$, and $P \geq 0$ are given.

The *receding horizon* control law applies

$$u(k) = \kappa(x(k)) := u_0^\star(x(k)),$$

where $u^\star(x(k))$ minimises $V(x(k), u)$.

(a) State clearly the receding horizon principle. Assume that K is any feedback gain such that $\rho(A + BK) < 1$ and that $P > 0$ satisfies

$$(A + BK)^\top P(A + BK) - P \;\leq\; -Q - K^\top RK.$$

Using the "shifted sequence + terminal tail" argument, show that the optimal value function $V^\star(x) = \min_u V(x, u)$ satisfies

$$V^\star(x(k+1)) \leq V^\star(x(k)) - x(k)^\top Q x(k).$$

Hence explain why the origin is an asymptotically stable equilibrium point of the *unconstrained* receding-horizon closed loop system. [25%]

(b) Now consider the *constrained* MPC problem with mixed linear constraints

$$Mx_i + Eu_i \leq b, \quad i = 0, 1, \ldots, N-1,$$

 (cont.

and a terminal constraint

$$M_N x_N \le b_N.$$

Show that the constrained MPC problem at time k can be written as a standard quadratic program (QP)

$$\min_{\theta} \ \frac{1}{2}\theta^\top H\theta + f^\top\theta \quad \text{s.t.} \quad A_{\text{eq}}\theta = b_{\text{eq}}, \ \ G\theta \le h,$$

using the stacked decision vector

$$\theta = \begin{bmatrix} u_0 \\ x_1 \\ u_1 \\ x_2 \\ \vdots \\ u_{N-1} \\ x_N \end{bmatrix}.$$

Give the *block structure* of H, A_{eq}, b_{eq}, G, and h in terms of $A, B, Q, R, P, M, E, b, M_N, b_N$ and the measured state $x(k)$. (You do *not* need to write every block entry explicitly, but you must show the matrix patterns clearly.) [35%]

(c) Even though the plant is linear, $\kappa(x) = u_0^\star(x)$ is generally *nonlinear* when constraints are present. Explain why. State what functional form $\kappa(x)$ takes when the optimisation is a strictly convex QP with linear constraints. [20%]

(d) Define:

(i) an *invariant set* for a closed-loop system $x^+ = f(x)$;

(ii) a *constraint-admissible* set for a feedback law $u = \kappa(x)$ with constraint set $\mathcal{Z} = \{(x, u) : Mx + Eu \le b\}$.

Explain how one can choose the terminal set $\{x : M_N x \le b_N\}$ to be a constraint-admissible invariant set for the linear tail law $u = Kx$, and how this choice guarantees *recursive feasibility* of the constrained receding-horizon controller. [20%]

4 Consider the discounted Markov decision process (MDP) with states

$$\mathcal{S} = \{s_0, s_1, s_2, s_T\},$$

where s_T is an absorbing terminal state. The action set is $\mathcal{A} = \{a, b\}$. The per-step *cost* is $c(s, a)$, and the discount factor is $\lambda = 0.9$.

The transition model is:

- From s_0:
  - action a: cost 1, next state s_1 with probability 1;
  - action b: cost 1, next state s_2 with probability 1.
- From s_1:
  - action a: cost 1, next state s_T with probability 0.9, and s_1 with probability 0.1;
  - action b: cost 2, next state s_2 with probability 1.
- From s_2:
  - action a: cost 1, next state s_T with probability 1;
  - action b: cost 2, next state s_0 with probability 1.
- From s_T: cost 0, remains in s_T.

(a) Define the value function $V^\pi(s)$ and action-value function $Q^\pi(s, a)$ for a deterministic policy π. Write down the Bellman equation satisfied by V^π. Write down the Bellman *optimality* equation for the optimal value function $V(s) = \min_\pi V^\pi(s)$. [15%]

(b) *Value iteration:* starting from $V_0(s) = 0$ for all states (and keeping $V_k(s_T) = 0$), compute V_1 and V_2 for s_0, s_1, s_2. From V_2, state the greedy policy at each non-terminal state. [25%]

(c) *Q-learning vs. SARSA (off-policy vs. on-policy):* two episodes are observed (each terminates upon reaching s_T):

- Episode 1: $(s_0, a, c = 1, s_1)$, $(s_1, a, c = 1, s_T)$;
- Episode 2: $(s_0, a, c = 1, s_1)$, $(s_1, a, c = 1, s_T)$.

Assume initial values $Q_0(s,a) = 0$ for all (s,a), learning rate $\alpha = 0.5$, discount $\lambda = 0.9$. Process the data in the order shown, updating after each transition.

(i) Using

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha\Big(c + \lambda \min_{b\in\mathcal{A}} Q(s',b)\Big),$$

compute the final Q-values after all transitions have been processed.

(ii) Using

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha\Big(c + \lambda Q(s',a')\Big),$$

where a' is the next action actually taken in the episode (as listed), compute the final Q-values.

(iii) Briefly explain why the two algorithms give different values here, and what this illustrates about off-policy vs. on-policy learning.

[30%]

(d) (i) Give two reasons why tabular $Q(s,a)$ becomes infeasible as the state dimension grows (or when s is continuous). [10%]

(ii) Consider a parameterised approximation $Q_\theta(x,a)$ and a target network with parameters θ_- held fixed for multiple updates. For a single sample (x,a,c,x'), define a squared TD loss

$$L(\theta) = \big(y - Q_\theta(x,a)\big)^2, \qquad y = c + \lambda \min_{b\in\mathcal{A}} Q_{\theta_-}(x',b).$$

Derive one stochastic gradient descent update for θ. [10%]

(iii) Explain the purpose of a replay buffer and a target network in stabilising learning. [10%]

END OF PAPER

THIS PAGE IS BLANK

Module 4F3: An optimisation based approach to control

(available in the exam)

1. (a) For the dynamical system satisfying, $\dot{x} = f(x, u),\ x(0) = x_0$, and the cost function

$$J(x_0, u(\cdot)) = \int_0^T c(x(t), u(t))\,dt\ + J_T(x(T))$$

then under suitable assumptions the value function, $V(x, t)$, satisfies the Hamilton-Jacobi-Bellman PDE,

$$-\frac{\partial V(x,t)}{\partial t} = \min_{u \in U}\left(c(x,u) + \frac{\partial V(x,t)}{\partial x} f(x,u)\right), \quad V(x,T) = J_T(x).$$

(b) For $f(x, u) = Ax + Bu$, $c(x, u) = x^T Q x + u^T R u$, and $J_T(x) = x^T X_T x$, if $X(t)$ satisfies the Riccati ODE,

$$-\dot{X} = Q + XA + A^T X - XBR^{-1}B^T X, \quad X(T) = X_T,$$

then $J_{opt} = x_0^T X(0) x_0$ and $u_{opt}(t) = -R^{-1} B^T X(t) x(t)$.

2. For the discrete-time system satisfying $x_{k+1} = Ax_k + Bu_k$ with x_0 given and cost function,

$$J(x_0, u_0, u_1, \ldots, u_{h-1}) = \sum_{k=0}^{h-1} \left(x_k^T Q x_k + u_k^T R u_k\right) + x_h^T X_h x_h,$$

if X_k satisfies the backward difference equation,

$$X_{k-1} = Q + A^T X_k A - A^T X_k B(R + B^T X_k B)^{-1} B^T X_k A,$$

then $J_{opt} = x_0^T X_0 x_0$ and optimal control signal, $u_k = -(R + B^T X_{k+1} B)^{-1} B^T X_{k+1} A x_k$.

3. For the system satisfying,

$$\begin{bmatrix} \dot{x} \\ \hline z \\ y \end{bmatrix} = \left[\begin{array}{c|cc} A & \begin{bmatrix} B_1 & 0 \end{bmatrix} & B_2 \\ \hline \begin{bmatrix} C_1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C_2 & \begin{bmatrix} 0 & I \end{bmatrix} & 0 \end{array}\right] \begin{bmatrix} x \\ \hline w \\ u \end{bmatrix} \quad \text{where} \quad \begin{cases} (A, B_2) & \text{controllable} \\ (A, C_1) & \text{observable} \\ (A, B_1) & \text{controllable} \\ (A, C_2) & \text{observable} \end{cases}$$

(a) The optimal $\mathcal{H}_2$ controller is given by,

$$\begin{bmatrix} \dot{x}_k \\ \hline u \end{bmatrix} = \left[\begin{array}{c|c} A - B_2F - HC_2 & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix}$$

where $F = B_2^T X$, $H = YC_2^T$, and X and Y are stabilising solutions to

$$0 = XA + A^T X + C_1^T C_1 - XB_2B_2^T X \quad \text{(CARE)}$$

and

$$0 = YA^T + AY + B_1B_1^T - YC_2^T C_2 Y \quad \text{(FARE)}$$

(b) The controller given by,

$$\begin{bmatrix} \dot{x}_k \\ \hline u \end{bmatrix} = \left[\begin{array}{c|c} \hat{A} - B_2F - HC_2 & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix}$$

where $F = B_2^T X$, $H = YC_2^T$, $\hat{A} = A + \frac{1}{\gamma^2} B_1 B_1^T X$, and X and Y are stabilising solutions to,

$$XA + A^T X + C_1^T C_1 - X(B_2B_2^T - \gamma^{-2} B_1 B_1^T)X = 0$$

and

$$Y\hat{A}^T + \hat{A}Y + B_1B_1^T - Y\big(C_2^T C_2 - \gamma^{-2} F^T F\big)Y = 0,$$

satisfies $\|T_{w\to z}\|_\infty \leq \gamma$.